



# Do the stimuli of an SSVEP-based BCI really have to be the same as the stimuli used for training it?

Andéol Évain, Ferran Argelaguet Sanz, Géry Casiez, Nicolas Roussel, Anatole Lécuyer

## ► To cite this version:

Andéol Évain, Ferran Argelaguet Sanz, Géry Casiez, Nicolas Roussel, Anatole Lécuyer. Do the stimuli of an SSVEP-based BCI really have to be the same as the stimuli used for training it?. *Brain-Computer Interfaces*, 2016, 3 (2), pp.103 - 111. 10.1080/2326263X.2016.1193458 . hal-01388534

**HAL Id: hal-01388534**

**<https://hal.inria.fr/hal-01388534>**

Submitted on 27 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Do the stimuli of an SSVEP-based BCI really have to be the same as the stimuli used for training it ?

Andéol Evain<sup>1</sup>, Ferran Argelaguet<sup>2</sup>, Géry Casiez<sup>3</sup>, Nicolas Roussel<sup>2</sup> and Anatole Lécuyer<sup>2</sup>

<sup>1</sup>Université de Rennes 1

<sup>2</sup>Inria

<sup>3</sup>Université de Lille

Friday 20<sup>th</sup> May, 2016

## Abstract

Does the stimulation used during the training on an SSVEP-based BCI have to be similar from the one of the end use ? We recorded 6-channels EEG data from 12 subjects in various conditions of distance between targets, and of difference in color between targets. Our analysis revealed that the stimulation configuration used for training which leads to the best classification accuracy is not always the one which is closest to the end use configuration. We found that the distance between targets during training is of little influence if the end use targets are close to each other, but that training at far distance can lead to a better accuracy for far distance end use ( $p < 0.01$ ). Additionally, an interaction effect is observed between training and testing color ( $p < 0.001$ ) : while training with monochrome targets leads to good performance only when the test context involves monochrome targets as well, a classifier trained on colored targets can be efficient both for colored and monochrome targets. In a nutshell, in the context of SSVEP-based BCI, training using distant targets of different colors seems to lead to the best and more robust performance in all end use contexts.

**Keywords:** BCI, SSVEP, visual stimulation, stimulus, training conditions, color, distance

## 1 Introduction

In a great number of BCI systems, a stimulation is presented to the user in order to observe and classify his cerebral response. This is notably the case for P300 or SSVEP-based BCI, which are sometimes called "exogenous" BCI [19], relying on external flashing or flickering stimuli. In such BCI setups, the properties of the stimulation must be carefully considered. In the case of SSVEP for example, one must choose: the size of the flickering targets, the distance between them, their flickering frequencies, and even their color or difference in color (using different colors for different targets). Target size and frequencies, as well as inter-target distance, were found to strongly influence the resulting classification accuracy of the BCI [5, 20, 23].

In recent years, signal processing methods for SSVEP-based BCIs that do not require a calibration phase have been proposed, and were shown to have a good accuracy, compared to the existing alternatives [9, 24]. However, an interest toward the capacity of calibration-based methods to take into account subject specificity was maintained, and several recent studies still use calibration-based methods [22, 15, 27].

In some cases, the characteristics of the stimulation vary between the training and the testing phases. SSVEP-based BCIs are sometimes trained in stimulation conditions which differ from the ones on the online use. As an example, in [16], only one flickering target at a time is presented to the user during

the training phase, while two different targets are presented during the test phase of the experiment. On top of that, the targets used for training are much larger than the online ones. In other cases, the properties of the stimulation are kept rigorously identical during both phases. For example, in [18], the flickering targets displayed during the training phase are identical to the targets used during the online BCI-based video-game. But very little is actually known regarding the potential influence of a difference in stimulation conditions between training and end use of exogenous BCI. In other words, is it preferable to train a BCI classifier in a stimulation context which is close to the end use context, i.e, the one in which the classifier will be used later?

It has already been observed that some SSVEP stimulation conditions, (thereafter referred to as "*easy condition*") could lead to a better classification accuracy [5, 20] (such as a larger distance between targets), whereas other stimulation characteristics (thereafter referred to as "*hard condition*") could lead to worse performance (resp. close distance). Therefore, is it preferable to train the BCI classifier in the *easy condition* or in the *hard condition*? In other words, is it better to "learn the easy way" or to "learn the hard way"? Limiting noise in the data with the *easy condition* could result in a better classifier. But on the other side, training the system in the *hard condition* could result in a more robust classifier, as "he who can do more, can do less". Besides, in the Machine Learning community, it

is generally considered that the training has to be done in conditions as close as possible to the end use context (from now on called "*end use condition*"). One last possibility is that mixing different conditions during the training, with both *hard condition* and *easy condition* trials, would result in a more robust classifier.

Thus, in this study, we explore which stimulation condition enables to train the most robust classifier in an SSVEP context. To sum up, four strategies for the optimal training conditions are possible: the *easy condition*, the *hard condition*, the *end use condition*, and the *mix condition*. We focused on two parameters of SSVEP visual stimulation : distance between targets, and difference in color between targets. Distance between targets is known to influence the accuracy of classification [20]. With a short distance between targets, SSVEP responses for several frequencies are observable in cerebral activity, thus making classification harder [20]. Visual selective attention is known to be affected by color, and this effect can be observed in the measurable cerebral activity [13]. Thus, it is plausible that using targets of different colors might have an effect on SSVEP response.

## 2 Related work

SSVEP detection is largely used for BCI applications which allow one user to select a target in a defined set [16, 17]. SSVEP requires an external visual stimulation of the user. Different stimuli have been used for SSVEP BCI-based applications, as reviewed by [28]. This meta-analysis showed that the most frequently used stimuli could be divided in three categories. The first one is graphic alternation, in which a simple form (circle, rectangles) is successively displayed and hidden at the preferred frequency. The second one is flickering checkerboards, which can be seen as a specific case of the first category, with checkerboard as the shape of the pattern. Finally, the third one consists in using light devices such as LEDs, allowing to control precisely the luminosity and the frequency of stimulation, but requiring specific hardware. LED-based stimulation allows for higher bit rate, higher stimulation frequencies, and more variability in the range of frequencies, while standard screens are limited by their refresh rate [28]. On average, graphic alternation on the screen results in a higher bit rate than flickering checkerboards. The same review also revealed that stimulation frequencies in the low (1-12 Hz) and medium (12-30 Hz) frequency band have been more often applied than those in the high frequency band (30-60 Hz). This tendency is probably due to both the limitations in frequency of computer screens, and the limited SSVEP response of the human visual cortex to such frequencies. Other studies [12, 21] found that frequencies below 10 Hz and above 25 Hz perform significantly less well than intermediary ones.

The authors concluded that : "improvement to stimuli can enhance SSVEP SNR (*signal-to-noise ratio*), simplify signal processing, enable the use of more targets, prevent loss of attention, and allow for BCI independent BCI operation".

However, little is known about the influence of stimulus color for SSVEP. Cao *et al.*[5] tried different colors flickering on a black background. They observed that white stimuli lead to the highest performance, followed by gray, red, green and blue. These colors are traditionally used for SSVEP experiments, and correspond to the RGB system basic colors, with the addition of a middle gray. Since then, other studies have analyzed the influence of color on SSVEP BCIs [1, 25, 26, 7]. Yet, existing studies consider that at any given time, all SSVEP targets share the same color both in training and in end use conditions. In contrast, in this paper, we introduce the usage of different colors for simultaneous SSVEP targets, each target being assigned a different color.

Other parameters, such as frequency, spatial size, number of simultaneously displayed stimuli, and their spatial proximity have been studied by Kian *et al.*[20]. They concluded that optimal performance is attained when targets are distant of at least 0.087 rad apart, with a size that subtends 0.035 rad of visual angle. Additionally, other studies [12, 21] found that low (below 10Hz) and high frequencies (above 25 Hz) perform significantly less well than intermediary ones.

In this paper, we study the effect of distance and difference in color between targets, not only for the end use, but also during the training phase, and the link between optimal training conditions and *end use conditions*. In particular, we focus on the influence of distance, and of difference in color between targets.

## 3 Materials and method

In order to explore the effect of distance and difference in color between targets, we registered SSVEP data in different conditions, considering two factors. The first factor is the distance between targets, which can take three different values. The second factor is the difference in color between targets. In the "monochrome" condition, targets are all black with a white background, while in the "color" condition, they each have a different color, with the same white background. For each of the conditions of distance and color, one data set was recorded. This data set was partitioned into 6 folds. For each fold, a classifier was trained on 5/6<sup>th</sup> of the original data set. The resulting classifier was then tested on the last 6<sup>th</sup>, and on the data from all the other conditions. For each condition, the rate of correct classification is measured by averaging over the folds. We could then study the variations of accuracy of a classifier, depending on both the training conditions and the testing conditions.

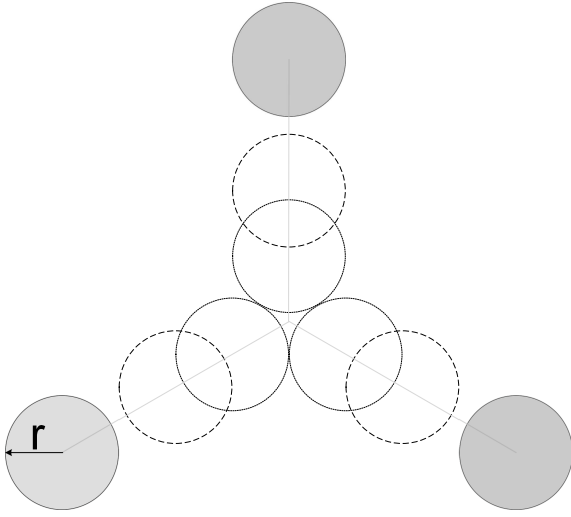


Figure 1: Different distances between targets (three possible distances :  $2r$ ,  $4r$ , and  $8r$ ).

**Participants:** 12 healthy participants volunteered to participate in the study : 10 men and 2 women, aged between 23 and 37 years old (average 28 years old, standard deviation 3.9 years). All of them happened to be right-handed, and had a normal or corrected vision. Before the experiment, each participant signed a written consent form, and filled a questionnaire collecting statistics about gender, dominant hand, age and sight. One participant reported not having followed the instructions. Thus, another participant passed the experiment to compensate for the discarded data.

**Data acquisition:** EEG data were recorded using 6 passive electrodes out of a 16-channel system (g.USBamp, g.tec company, Austria), with a sampling rate of 512 Hz. Electrodes were placed according to the extended 10-20 system [14], concentrated on the occipital lobe, in order to focus on the visual cortical areas. Electrodes were positioned on CPz, POz, Oz, Iz, O1 and O2. A reference electrode was located on the right ear, and an additional ground electrode was located on AFz. Channels were amplified and band-pass filtered between 2 and 60Hz. A notch filter was applied to exclude frequencies between 48 and 52Hz, corresponding to the power supply frequency band. Electrode impedance was checked to be below 1 kilo-ohm to ensure signal quality.

**Design:** Inter-stimuli distances and color conditions varied across the experiment, with 3 levels of distance (See Figure 2), and 2 levels of color. All the other characteristics of the stimuli were kept constant.

On a 60 Hz screen, it is possible to display frequencies at 8.57 Hz or below, and at 10, 12, 15, 20, 30 or 60 Hz. Frequencies lower than 10 Hz and above 25 Hz were excluded from this study because of their weaker

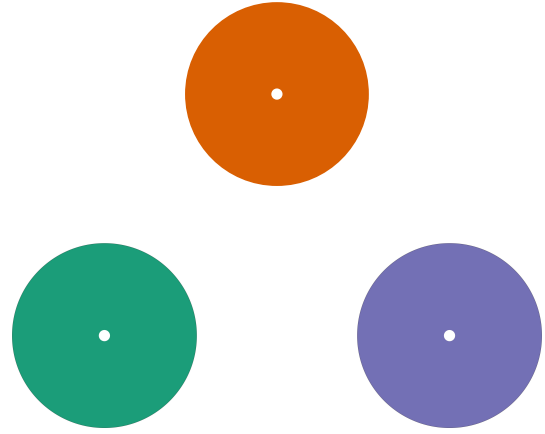


Figure 2: Targets shape, and their respective colors in the colored condition: left target is green, right target is blue, and upper target is orange.

SSVEP response [12, 21, 28]. Finally, SSVEP brain response is observable at the harmonics of the stimulation frequencies. Thus, a 20 Hz frequency could interfere with the detection of a 10 Hz one. We chose to use 3 targets flickering respectively at 10, 12 and 15 Hz. The screen vertical synchronization was used to control each frequency.

In order to reach an optimal detection, stimuli should be at least 0.035 rad wide [20]. On top of that, a large inter-target distance is preferable, until 0.087 rad [20]. The targets had to be large enough to allow a good target detection, but small enough to be displayed close to each other without overlapping. We chose to use targets with a diameter of 0.05 rad of visual angle. Considering that the participants would be sitting between 50 and 70 cm of the screen, 0.05 rad of visual angle corresponds to 3.1 cm on the screen. We used a DELL™Ultrasharp™2007FP 51 cm screen (20.1 inches), with a resolution of  $1280 \times 1024$  pixels, and a refresh rate of 60 Hz, resulting in a diameter of 58 pixels (3.1 cm) for each stimulus.

The distance between targets varied during the experiment, with 3 possible distances: In the "close" condition of distance, the targets were touching each other, without overlapping (3.1 cm between the centers of 2 stimuli, exactly 58 pixels). In the "medium" condition of distance, the targets were twice as far, at 6.3 cm of each other. Finally, in the "far" condition, the distance is again doubled, reaching 12.6 cm between the centers of each target (See Figure 2).

In the "monochrome" condition, the stimuli were flickering between pure black and white (RGB: 0, 0, 0 and RGB: 255, 255, 255), one at 10 Hz, one at 12 Hz, and the last one at 15 Hz. In the "colored" condition, the frequencies of the stimuli were the same, but they were each associated with a specific color, and flickered between white and this color. In order to exhibit a possible effect of a difference in color between targets, colors were chosen to be as different as possible from



each other, in the sense of attention selection. The chosen colors should also present a good contrast with the white background (as luminosity contrast is known to be essential to SSVEP responses [6]), that is, a similar level of gray. The GeoVista center of Penn State University proposed several graphic palettes specially designed for attention differentiation [4]. Among those palettes, the final choice of colors was the one presenting the highest level of gray, and guaranteed to be colorblind-friendly [10]. It was composed of a green (RGB: 27, 158, 119), an orange (RGB: 217, 95, 2), and a purple (RGB: 117, 112, 179) (see Figure 1). Some interactions between color and frequency of stimulation have been discovered [5, 23], but they are not well understood yet. In this paper, we chose not to focus on those aspects, in order to maintain a low number of factors, and thus keep a good statistical power. Therefore, we associated each frequency to a color, and kept it constant. This association was chosen in order to get the strongest SSVEP response, according to the present knowledge. The SSVEP response to a red stimulation is stronger at 12 Hz [5, 23]. Thus, we associated the 12 Hz target with the color with the strongest red component (orange). Others colors have not been shown to have such a strong dependency toward the frequency of stimulation for SSVEP response [23]. The green was associated with the 10 Hz target, and the purple with the 15 Hz one.

**Experimental procedure:** Participants were seated in a comfortable chair, in front of a computer screen, on which flickering stimuli were displayed, while their cerebral activity was recorded. The experiment was divided into eight blocks, separated by one minute breaks.

Before each trial, instructions, telling the participants which of the 3 targets they had to focus on, were displayed on screen for 3 seconds. The targets then flickered during 7 seconds, and a simulated feedback on the selected target was then displayed.

In order to keep the attention of the participants, it is important to give them a feedback indicating what target has been recognized [17]. However, a training phase for the feedback could induce a bias toward the condition of feedback training. In order to avoid an augmentation of the number of conditions, a fake feedback was used. It has been shown that even a fake feedback can help to keep the attention of a participant during BCI tasks [11]. Thus, after each trial, a fake feedback was displayed during 2 seconds, giving a positive response with a realistic accuracy (80%). This level of accuracy has been reached in pre-experiments, as well as in previous studies on SSVEP [2, 17]. The users were not aware that the feedback was simulated.

Each block of acquisition contained 18 trials, and lasted 3 minutes and 36 seconds. The whole experiment was composed of 8 blocks. Between each

block, participants were given at least a one minute break that could be extended upon participants will. All participants did all of the blocks (balanced design), with randomized order of conditions. The distance between the targets and the colors of the targets changed across blocks. Finally, the participants were given a post-experiment questionnaire. In particular, this questionnaire allowed us to check if the participants trusted the feedback. The total duration of the experiment was less than 45 minutes.

Each block was characterized by the distance between the targets  $D$ , and the color of the targets  $C$ . There were 3 distance levels: close ( $D_c$ ), medium ( $D_m$ ) and far ( $D_f$ ), and 2 color levels: monochrome ( $C_m$ ) and colored targets ( $C_c$ ). Each of the 6 resulting joint conditions of  $D/C$  corresponded to one block. On top of these 6 blocks, 2 additional blocks mixed all conditions with equal representation. The order of the final 8 blocks was randomized for each participant. In every block, each frequency was targeted 6 times. Each block was divided into 6 sub-blocks in which every frequency was targeted once, in order to avoid any bias in the frequency distribution across the block. In the blocks that mixed the conditions, each combination of target, color and distance was presented exactly one time.

In summary, the experimental design was 12 participants  $\times$  8 blocks (3 distances  $\times$  2 colors + 2 mixed)  $\times$  18 trials = 1,728 total trials.

## 4 Signal processing

Most SSVEP studies use similar steps of filtering, feature extraction, and classification. The signal processing chain used for this study followed classical methods for each step [22].

**Feature extraction:** A measure of the spectral density in each frequency of interest is computed as in [18]. Each channel was processed through a 0.5 Hz-wide band-pass filter at the 3 fundamental frequencies of stimulation (10, 12 and 15 Hz), and their first harmonic (respectively 20, 24 and 30 Hz). For each frequency of stimulation, the six channels are processed through two fourth order common spatial pattern (CSP) filters, in order to optimize the detection of this specific frequency [18]. The resulting filtered signals are then decomposed in 0.5 seconds moving windows, with 0.1 seconds moving steps. Let  $S(f)$  be the signal filtered around frequency  $f$ . The energy spectral density is computed as the average of  $S^2(f)$  over the time window, and a natural logarithm of this estimated density is computed and used as feature for the following classification algorithm.

**Classification:** A three-class LDA classifier was trained, combining 3 two-class LDA classifiers, each of them discriminating one class versus all the others. For each class  $i$  of stimulation (defined by its frequency), a two-class LDA classifier is learned, discriminating signals of class  $i$  against all the others. This classifier gives  $d_i$  an oriented distance to the hyperplane of separation. When used with two classes, such a classifier decides for class  $i$  if  $0 \leq d_i$ . In order to combine several two-class LDA classifier to classify between more classes, the chosen class is the one maximizing  $d_i - \sum_{j \neq i} d_j$ . The resulting 3-class classifier gives a decision every 0.1 seconds, and thus has a relatively low recognition rate (71.8% on average during this experiment), since it makes a decision using only 0.5 seconds of signal. In order to get one single classification for each 7-seconds trials, a voting step could be added, by deciding for the class the most consistently detected over the 7 seconds. Such a voting system allows reaching recognition rates close to 100% on 7-seconds trial. The resulting accuracies are strongly discrete, since the dataset were only 18 trials long. Thus, to avoid discretisation of measurements, the following performance analysis uses the accuracy measured on 0.5-seconds windows.

## 5 Data analysis and results

The following analysis focuses on the effect of the difference in color and distance of the SSVEP stimuli on the classification performance. For each pair of blocks (characterized by its level of Distance and Color), we will refer to it as training block when its data is used to train the classifier and as test block when it is used to test the classifier. The matrix of performance  $P$  provides for each cell  $(x,y)$  (row,column) the average classification rate for the classifier trained using the  $x$  training block, and tested on the  $y$  test block (See Table 1). The accuracy for each cell has been computed by 6-fold trial-based cross-validation. For each block, the dataset is divided into 6 folds. The folding does not separate data from the same trial, in order to avoid overlapping in samples. It also ensures that classes are equally represented between folds. For each fold,  $5/6^{\text{th}}$  of the dataset is used to train a classifier. This classifier is tested on the last  $6^{\text{th}}$  (to get accuracy when  $y = x$ ), and on the dataset from the other conditions (to get accuracy when  $y \neq x$ ). Accuracies are then averaged over the folds. This method allows to compute the diagonal of  $P$ , without introducing a bias due to the length of the training set, and without recording each condition twice, which would have made the experiment twice as long, and would not have been acceptable for the participants, especially since SSVEP stimulation causes a strong visual fatigue.

Before the analysis, let us first introduce the notation used. Each block was characterized by the distance  $D$  between targets (close ( $D_c$ ), medium ( $D_m$ )

and far ( $D_f$ )) and their color  $C$  (monochrome ( $C_m$ ) and colored targets ( $C_c$ )). We write each condition as "Distance condition/Color condition". For example,  $D/C_c$  represents the set of all blocks with colored targets, while  $D_m/C_c$  only represents the block with medium distance and colored targets. Mixed blocks were noted as  $Mix_1$  and  $Mix_2$  ( $Mix$  represents both blocks).

Regarding the overall performance rates, classification accuracy was 71.8%. As said in the previous section, this accuracy is measured on 0.5 seconds sliding windows for each trial (7 seconds).

### 5.1 Single conditions analysis

We performed a four-way ANOVA with a full factorial design considering the training ( $x = \{D/C\}$ ) and testing ( $y = \{D/C\}$ ) conditions. When needed, Tukey post-hoc pairwise-tests were performed ( $\alpha=0.95$ ). As the  $Mix$  condition can not be attributed a level of distance or color, it was considered in a second analysis.

The four-way ANOVA only showed a significant main effect on test distance ( $F(2, 22)=5.3$ ;  $p<0.05$ ;  $\eta_p^2=0.325$ ). Post-hoc tests showed that  $y = \{D_c/C\}$  resulted in a significantly lower accuracy rate ( $\bar{x}=69.8$ ;  $\sigma=14.7$ ) than  $y = \{D_m/C\}$  ( $\bar{x}=72.5$ ;  $\sigma=14.3$ ).

Regarding higher interaction effects, there was a significant two-way interaction effect between training and testing distance conditions ( $F(4, 44)=5.9$ ;  $p<0.01$ ;  $\eta_p^2=0.349$ ), as well as between training and testing color conditions ( $F(1, 11)=26.63$ ;  $p<0.001$ ;  $\eta_p^2=0.71$ ) (see Figure 3). The interaction between training and testing distance conditions can be interpreted as that training with ( $x = \{D_c/C\}$ ) provides similar performances for all ( $y = \{D/C\}$ ) while training with ( $x = \{D_m/C\}$ ) and ( $x = \{D_f/C\}$ ) results in similar performances for ( $y = \{D_c/C\}$ ) but increased performances with ( $y = \{D_m/C\}$ ) and ( $y = \{D_f/C\}$ ). Regarding training and testing color, post-hoc tests showed that training the classifier with monochrome targets ( $x = \{D/C_m\}$ ) resulted in a significantly lower classification accuracy for colored targets ( $y = \{D/C_c\}$ ). On average, classifiers trained on monochrome targets presented an accuracy 8.6% lower when tested on colored targets.

We also analyzed the effect of training and testing the classifier with the same conditions ( $y = x$ ) and the opposite case ( $y \neq x$ ). We performed a two-way ANOVA dataset ( $(y = x)$ ,  $(y \neq x)$ ) and training condition vs accuracy. The results show that the accuracy of a classifier trained in the same condition as the end use ( $y = x$ ) have a significantly better average accuracy than the ones which are not ( $y \neq x$ ) ( $F(1, 11)=43.50$ ;  $p<0.001$ ;  $\eta_p^2=0.79$ ). However, this effect is observed only when comparing the diagonal with all the training options at once, and does not re-

condition	$D_c/C_m$	$D_m/C_m$	$D_f/C_m$	$D_c/C_c$	$D_m/C_c$	$D_f/C_c$	Mix1	Mix2
$D_c/C_m$	76.0	73.7	73.3	61.7	66.1	63.3	71.1	68.3
$D_m/C_m$	72.1	76.7	75.1	64.3	68.4	66.7	74.1	68.3
$D_f/C_m$	71.9	76.8	75.8	66.1	69.0	68.4	73.5	70.6
$D_c/C_c$	70.7	72.2	73.8	70.6	70.4	70.7	74.4	70.9
$D_m/C_c$	70.6	75.4	74.9	69.6	74.6	71.7	76.4	72.7
$D_f/C_c$	73.0	74.2	75.0	71.2	73.1	74.0	75.2	72.1
Mix1	72.9	76.3	75.2	68.5	71.3	69.5	76.3	71.7
Mix2	72.8	75.5	75.7	68.0	70.9	69.4	74.7	71.4

Table 1: Performance matrix averaged over the 12 participants.  $D_c$ ,  $D_m$  and  $D_f$  are respectively the close, medium, and far distance levels.  $C_m$  and  $C_c$  are the monochrome and colored levels of color. Each cell contains the accuracy obtained with a classifier trained in the condition specified by the line, tested in the condition specified by the column.

main significant when focusing on more specific conditions.

In particular, when comparing with the classifier trained on colored targets ( $x = D/C_c$ ), we found no significant differences in accuracy. These results suggest that training a classifier in the end use condition gives good results, but that similar results could be obtained with another training condition (here with a classifier trained on colored targets at medium or far distance) and, as it will be presented later in section 6, such other classifier can be used for a wide range of end use conditions, meaning without the need for re-computing a classifier after every change in the end use condition.

## 5.2 Mix condition analysis

We tested the quality of the classifier when trained with the *Mix condition*  $x=\{Mix\}$  and tested against the different combinations of color and distance  $y=\{D/C\}$ .

The two-way ANOVA test distance and test color vs accuracy showed a main effect on distance ( $F(2, 22)=6.19; p<0.01; \eta_p^2=0.36$ ). Post-hoc tests showed that the mean accuracy of  $D_c$  ( $\bar{x}=70.55; \sigma=16.4$ ) was significantly smaller than for  $D_m$  ( $\bar{x}=73.5; \sigma=15.22$ ). We also observed a tendency on test color ( $F(1, 11)=5.9; p=0.08; \eta_p^2=0.249$ ) which suggests that  $y=\{D/C_m\}$  would result in higher classification rates.

In addition, we tested the accuracy of the classifier when training it with the different "single" conditions  $x=\{D/C\}$  and tested against the *Mix conditions*  $y=\{Mix\}$ . The two-way ANOVA training distance and training color vs accuracy showed a strong tendency on training color ( $F(1, 11)=4.62; p=0.055; \eta_p^2=0.296$ ) and a tendency on training distance ( $F(2, 22)=2.88; p=0.078; \eta_p^2=0.207$ ). This results are in line with the results of the single condition analysis as the tendencies suggest that ( $x = \{D/C_c\}$ ) would result in higher classification rates than ( $x = \{D/C_m\}$ ), and ( $x = \{D_c/C\}$ ) would result in lower classification rates than ( $x = \{D_m/C\}$ ) and ( $x = \{D_f/C\}$ ).

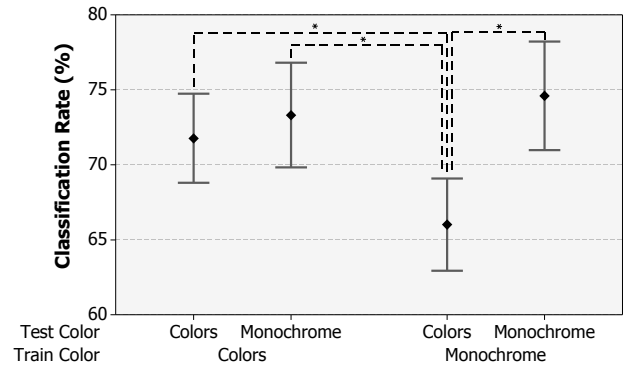


Figure 3: Classifier accuracy (CI 95%) depending on the training and testing color. Left: classifiers trained with colored targets tested on colored and monochrome targets. Right: classifiers trained with monochrome targets tested on colored and monochrome targets.

In short, in order to get a general-purpose SSVEP classifier able to adapt to various conditions, training a classifier on distant targets is more efficient than training them on short distance. This effect could be explained by the noise generated by the interference between targets. We also saw that training a classifier on colored targets seems to be more efficient, when the end use presents colored targets, and as efficient when the end use presents monochrome targets.

## 5.3 Subjective evaluation

Participants filled a questionnaire aiming at evaluating perceptual differences between the conditions of stimulation.

Concerning the detection of the fake feedback, a 7-point Likert-scale was used to rate the precision of the target detection and the control of the target detection. For all participants, with all conditions, the feedback had 80% accuracy. The precision of the target detection was rated 5.8 on average ( $SD = 0.4$ ), and control of the detection was rated 5.3 ( $SD = 1.2$ ). This high

level of control of the detection indicates that most subjects did not suspect the fake feedback. One participant rated 5 for the detection and 2 for the control. In addition, this participant stated that *"the errors are surprising: sometimes no mistake while I had the impression of having a bad focus, and when there is a mistake, the detected target is not the one that was distracting me"*. This participant could have discovered the fake feedback. However, his data resulted in a good accuracy (71% of correct classification on average), without any drop of performance on specific blocks, as it should be the case for someone trying to not follow the instruction. The other participants all rated control of the detection at 5 or higher.

Concerning subjective preference between color and monochrome targets, opinions differs across the subject. They were invited to report any impression with their own words. Some participants preferred the colored targets, while others preferred the monochrome ones. For example, one participant stated : *"About black and white : I don't know if the precision is actually better, but it is undeniably more comfortable to use"*, while another one said the opposite : *"Overall, the flickering in black and white targets seemed more annoying to bear with"*. Most participants did not comment about the distance between targets. This suggests that the distance between targets did not have a meaningful influence on user comfort.

To sum up with, difference in color between targets seems to have contrasted effects depending on the participants, while the distance between targets has little influence on comfort. This observation should encourage future applications to use distant targets for both training phase and online mode whenever it is possible, since the resulting classification rate is higher. As for color, considering that using colored target during end use results in a similar classification rate, and that the comfort related to color seems to be participant-dependent, the choice of using colored or monochrome targets during the end use could be left to the user.

## 5.4 Results summary

We observed an effect of the distance between targets for the testing condition. When SSVEP targets are closer to each other, they are harder to classify. This result is consistent with previous research [20]. Additionally, an interaction effect was observed between training distance condition and test distance condition. It can be interpreted as follows: A classifier trained at close distance is equally accurate for targets at close or far distance, while a classifier trained with targets at far distance is more accurate for an end use at far distance, and equally accurate at close distance. In short, the distance condition seems to follow a rule of "learning the easy way".

The color condition behaves quite differently. No significant effect of color on accuracy was found neither during the training phase, nor during the end use. However, an interaction effect was observed between the training color condition and the testing color condition. It can be interpreted as follows: a classifier trained with monochrome targets is efficient for an end use with monochrome targets, but is less accurate in the presence of colored targets. On the opposite, a classifier trained with targets of different colors is accurate during the end use for both monochrome and colored targets. In short, the color condition seems to follow a rule of "learning the hard way".

The main lesson, here, is that a classifier trained in the right conditions (medium or far distance, and different colors), can be used with a broader range of targets without any loss of accuracy compared to a specialized classifier trained on the exact same conditions as the end use.

## 6 Discussion

In machine learning, it is generally considered that it is preferable to train the classifier in conditions as close as possible to the end use ones. Since all parameters cannot be controlled, this general principle is most of the time restricted to the parameters considered to be the most important ones. For example, in an SSVEP context, the frequency of flickering is always conserved between training and end use. The core principle of SSVEP requires keeping such parameter constant. In addition, the targets color is generally conserved, as well as the distance between targets and the shape of those targets.

The same reasoning leads to mixing conditions during the training when the end use requires various conditions. However, our experiment suggests that this approach is not optimal, and that some conditions are more favorable for a robust training, even when extended to less favorable conditions.

We observed an interaction effect on the accuracy obtained depending on the color condition. In order to get a good accuracy when the end use context contains monochrome targets, both conditions of color during the training lead to similar performance. However, if the end context contains colored targets, training the classifier on monochrome targets becomes less efficient. Thus, if the end context is uncertain, we would recommend using targets of different colors for the training phase.

This effect can be explained, considering that adding color to the stimulation is equivalent to adding a noise to the data that does not change the class of each trial. Since the color does not change the frequency of stimulation on which the detection is based, the class actually stays unchanged. Thus, color might be a rele-



vant class invariant. Previous work in machine learning showed that adding a relevant noise to training data can improve the generalization power of a classifier [3, 8]. Understanding precisely how the neural structure of visual areas is related to this change when using targets of different colors still requires investigation. In particular, it would be interesting to explore whether the effects discovered in this study can be generalized to different sets and configurations of colors and frequencies of stimulation.

We observed an effect of the distance between targets during the end use. Targets separated by larger distances are easier to distinguish than targets close to each other. This is consistent with previous studies [20] that showed that when several targets flicker at different frequencies in the visual field of the user, all frequencies are represented in the cerebral activity, and compete for representation. The attention of the user then selects one frequency to improve its representation. Despite this attention selection, discrimination between close targets is still harder to achieve. Distant targets influence less the representation of the desired target, since the visual receptive cells are mainly condensed in the center of the visual field. Thus, for the distance parameter, the *easy condition* is when the targets are far from each other, while the *hard condition* corresponds to the closest distance.

We observed an interaction effect between training and testing distance conditions. Overall, an increase of the distance between targets during the test condition lead to an improved classification. However, this improved performance is less visible when the classifier is trained with closer targets. In contrast, a classifier trained with further away targets presents a stronger classification improvement between closer and further away targets. In short, it seems that training on the *easy condition* can be better than training a classifier on the *hard condition*. Concerning inter-stimulus distance, learning the easy way is more efficient than learning the hard way.

Let us now consider the *end use condition*: from our observations, in some cases, learning on conditions as close as possible to the *end use condition* is efficient. However, it does not seem to be always the case. For example, when the end use targets are close to each other, classifiers trained with targets at far distances are as efficient as the ones trained at short distances, with the additional advantage of being better for far distance targets. Similarly, classifiers trained with colored targets are as good as the ones trained on monochrome targets when it comes to classify monochrome targets, while classifiers trained on monochrome targets are less accurate to classify colored targets.

The comparison to the *Mix condition* reveals that the *mix condition* is not more efficient than the *far condition*, which corresponds to the *easy condition*. To sum up

with, the distance parameter follows the rule: learning the easy way.

Taken together, our results show that training a classifier in the end use condition is a relatively efficient approach. However, it is possible to choose the training conditions in order to ensure a robust accuracy for all end use conditions. Namely, training a classifier with targets at medium or far distance, and of different colors, leads to an accuracy similar to that of a training in the end use condition. In other words, by training a classifier in optimal conditions, the need for re-calibration when the end use conditions change is avoided. The classifier can be trained once and for all.

## 7 Conclusion

Similar conditions during the training of a classifier and its end use is not always better. Our findings show that it is possible to efficiently train a classifier on a specific condition and to generalize its use to different contexts. In some cases, it can even result in a better classification rate than when the classifier is trained in the end use context. In our SSVEP experiment, we found that distant targets are easier to distinguish, and are preferable as training conditions as well. Concerning the difference in color between targets, we found that targets of different colors are not significantly more difficult to distinguish. Last, we observed that using different colors for targets during the training phase results in a classifier efficient in a wider range of end use contexts.

Based on these observations, and when the end SSVEP application allows it, we advise to use distant targets with different colors during the training phase, and distant targets for the end use, while the end use color does not matter.

Future work could explore in more details the effect of conditions during the training and the end use. For BCI based on externally evoked potential, such as SSVEP and P300, on top of inter-target distance and difference in color, conditions such as raw color, frequency, or shape of the target could be relevant. Others parameters are relevant for internally evoked potential as well. A particular attention could be given to the instructions given to the users, or the feedback characteristics.

## References

- [1] Mustafa Aljshamee, Mahdi Q Mohammed, Abbas Malekpour Riaz-Ul-Ahsan Choudhury, and Peter Luksch. Beyond pure frequency and phases exploiting: Color influence in ssvep based on bci. *Computer Technology and Application*, 5(2), 2014.

- [2] Luzheng Bi, Ke Jie, Xin an Fan, and Yun Li. A SSVEP brain-computer interface with the hybrid stimuli of SSVEP and P300. In *ICME International Conference on Complex Medical Engineering (CME)*, pages 211–214, 2013.
- [3] Christopher M Bishop. Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7(1):108–116, 1995.
- [4] Cynthia A Brewer and Mark Harrower. Colorbrewer 2.0. <http://colorbrewer2.org/>.
- [5] Teng Cao, Feng Wan, Peng Un Mak, Pui-In Mak, Mang I Vai, and Yong Hu. Flashing color on the performance of ssvep-based brain-computer interfaces. In *IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1819–1822, 2012.
- [6] Galloway, NR. Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine. *The British journal of ophthalmology*, 74(4):255, 1990.
- [7] Vialatte, François-Benoît and Maurice, Monique and Dauwels, Justin and Cichocki, Andrzej. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology*, 90(4):418–438, 2010.
- [8] Zhao, Wenyi and Chellappa, Rama and Phillips, P Jonathon. *Subspace linear discriminant analysis for face recognition*. Citeseer, 1999.
- [9] Ola Friman, Ivan Volosyak, and Axel Gräser. Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE transactions on Biomedical Engineering*, 54(4):742–750, 2007.
- [10] Steven D Gardner. *Evaluation of the colorbrewer color schemes for accomodation of map readers with impaired color vision*. PhD thesis, Pennsylvania State University, College of Earth and Mineral Sciences, 2005.
- [11] Mar Gonzalez-Franco, Peng Yuan, Dan Zhang, Bo Hong, and Shangkai Gao. Motor imagery based brain-computer interface: A study of the effect of positive and negative feedback. In *IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6323–6326, 2011.
- [12] Christoph S Herrmann. Human eeg responses to 1-100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137(3-4):346–353, 2001.
- [13] Steven A Hillyard and Lourdes Anllo-Vento. Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95(3):781–787, 1998.
- [14] Herbert Henri Jasper. The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10:371–375, 1958.
- [15] Marcin Kolodziej, Andrzej Majkowski, and Remigiusz J Rak. A new method of spatial filters design for brain-computer interface based on steady state visually evoked potentials. In *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2015 IEEE 8th International Conference on*, volume 2, pages 697–700. IEEE, 2015.
- [16] Edmund C Lalor, Simon P Kelly, Ciaran Finucane, R Burke, Robert Smith, Richard B Reilly, and Gary McDarby. Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP Journal on Applied Signal Processing*, 2005.
- [17] Jozef Legény, Raquel Viciana-Abad, and Anatole Lécuyer. Navigating in virtual worlds using a self-paced SSVEP-based brain-computer interface with integrated stimulation and real-time feedback. *Presence*, 20(6):529–544, 2011.
- [18] Jozef Legeny, Raquel Viciana-Abad, and Anatole Lecuyer. Toward contextual SSVEP-based bci controller: Smart activation of stimuli and control weighting. *IEEE transactions on Computational Intelligence and AI in Games*, 5(2):111–116, 2013.
- [19] Mauro Marchetti, Francesco Piccione, Stefano Silvoni, and Konstantinos Priftis. Exogenous and endogenous orienting of visuospatial attention in p300-guided brain computer interfaces: A pilot study on healthy participants. *Clinical Neurophysiology*, 123(4):774 – 779, 2012.
- [20] Kian Bee NG, Andrew P Bradley, and Ross Cunningham. Stimulus specificity of a steady-state visual-evoked potential-based brain-computer interface. *Journal of Neural Engineering*, 9(3), 2012.
- [21] Maria A Pastor, Julio Artieda, Javier Arbizu, Miguel Valencia, and Jose C Masdeu. Human cerebral activation during steady-state visual-evoked responses. *The Journal of Neuroscience*, 23(37):11621–11627, 2003.
- [22] Liu Quan, Chen Kun, Ai Qingsong, and Sheng Quan Xie. Review : Recent development of signal processing algorithms for ssvep-based brain computer interfaces. *Journal of biomedical and Biological Engineering*, 34:299–309, 2013.
- [23] David Regan. An effect of stimulus colour on average steady-state potentials evoked in man. *Nature*, 210:1056–1057, 1966.
- [24] Kuo-Kai Shyu, Yun-Jen Chiu, Po-Lei Lee, Ming-Huan Lee, Jyun-Jie Sie, Chi-Hsun Wu, Yu-Te Wu, and Pi-Cheng Tung. Total design of an fpga-based brain-computer interface control hospital bed nursing system. *IEEE Transactions on Industrial Electronics*, 60(7):2731–2739, 2013.

- [25] Rajesh Singla, Arun Khosla, and Rameshwar Jha. Influence of stimuli color on steady-state visual evoked potentials based bci wheelchair control. *Journal of Biomedical Science and Engineering*, 6(11):1050, 2013.
- [26] Richard Junior Manuel Godinez Tello, Sandra Mara Torres Müller, André Ferreira, and Teodiano Freire Bastos. Comparison of the influence of stimuli color on steady-state visual evoked potentials. *Research on Biomedical Engineering*, 31(3):218–231, 2015.
- [27] Yuliyana Velchev, Dimitar Radev, and Svetla Radeva. Features extraction based on subspace methods with application to ssvep bci. *International Journal*, 52, 2016.
- [28] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M. Aarts. A survey of stimulation methods used in ssvep-based bcis. *Intell. Neuroscience*, 2010:1:1–1:12, 2010.